DOCUMENT RESUME

ED 399 771 FL 024 096

AUTHOR Moritz, Christine E. B.
TITLE Student Self-Assessment of Language Proficiency: Perceptions of Self and Others.
PUB DATE 24 Mar 96
NOTE 22p.; Paper presented at the Annual Meeting of the American Association for Applied Linguistics (18th, Chicago, IL, March 23-26, 1996).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Processes; College Students; Higher Education; Interviews; Language Aptitude; *Language Proficiency; Language Research; Protocol Analysis; Questionnaires; *Second Languages; Self Concept; *Self Esteem; *Self Evaluation (Individuals)

ABSTRACT

A study explored the cognitive processes and social-situational influences underlying students' assessment of their own foreign language proficiency, focusing on process rather than result of self-evaluation. The subjects, 28 college student learners of French at different course levels, were administered a self-assessment questionnaire on the four language skill areas (listening, speaking, reading, writing). Subsequently, two types of verbal report were elicited: a think-aloud protocol and an immediate, semi-structured retrospective interview. Data were analyzed for evidence of: (1) student orientation to the self-assessment task, (2) interpretation of questions and rating scales, (3) possible influence of course level and previous language experience, and (4) students' basic level of comfort in speaking French. Six categories of factors influencing the self-assessment process were identified (question interpretation, language learning background/experience, reference points, questionnaire-completion strategies, level of certainty about answers, self-confidence level). Results show students use a variety of reference points/benchmarks when evaluating their own language abilities, particularly social category, meaningful other, and autobiographical. Pedagogical implications are discussed briefly. Contains 23 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 399 771

Student Self-Assessment of Language Proficiency: Perceptions of Self and Others

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Christine E.B. Moritz
The University of South Carolina

a paper presented at
AAAL Chicago -- March 24, 1996

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Christine E. B Moritz
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Introduction:

Self-assessment in foreign language education has become a relatively popular issue of discussion and research in the past few decades (see Bachman & Palmer 1989, Blanche & Merino, 1989, Heilenman, 1990 & 1991, LeBlanc & Painchaud, 1985, Oskarsson, 1988 & 1989, and Wesche, Parikbaht and Ready, 1993, to name just a few). This attention is due in part to a growing interest in non-traditional forms of assessment, in part to the appeal of self-assessment as a logical component of learner-centered pedagogy, and in part to its alleged potential to alleviate the testing workload of teachers. In several European countries, self-assessment is used quite extensively, often in language programs for adult immigrants, and almost always as a **formative** testing instrument, that is, as a means by which students may continuously monitor and evaluate their progress over the course of their language learning experience. (see for example, Holec, 1980, 1988; Oskarsson, 1984, among others). In fact, self-assessment has its roots in the movement for more self-directed (autonomous) learning programs. In North America, there has been relatively little discussion of self-assessment's promise as a formative pedagogical tool; attention has concentrated instead on its **summative** capacities, that is, as a means of making judgments about learning **after** it has taken place. In particular, there has been

FL024096

BEST COPY AVAILABLE

a good deal of interest in the notion that self-assessment can be used as an effective method of placement testing (see for example, discussions in Heilenman, 1991, and LeBlanc & Painchaud, 1985).

Enthusiasm for self-assessment and its apparent advantages, has to a certain extent, overshadowed the question of what exactly is involved when students are asked to describe themselves according to a pre-determined scale. Though many have claimed that self-assessment is an effective measurement of second language proficiency, only a few have addressed the question of what factors other than knowledge of language proficiency influence students' self assessments. This paper describes one aspect of a study on the social and psychological variables that may affect how second language learners orient themselves to a self-assessment task. The results of this research show that students' perceptions of their language abilities are complex, and shaped by many factors. I contend, therefore, that there is good reason to seriously question whether or not we can reasonably apply this methodology in a placement testing situation, which entails a comparison between individuals' self-assessments.

Brief overview of Higher Education and Foreign Language Research:

Before turning to a description of the present study, I'd like to present a very brief overview of some of the previous research on self-assessment in Foreign Language education. Between 1970 and 1995, there were 53 publications concerning self-assessment in foreign language education. Of these, 39 described empirical studies. A careful review of this research has exposed three significant recurring gaps and problems. First, the majority of articles reporting on self-assessment in foreign language learning make only slight (if any) reference to the very large body of research that has been reported concerning self-assessment in other fields

of higher education (a search of the ERIC database alone generated over 400 articles specifically addressing the use of self-assessment in higher education). This general education literature acknowledges that there is **no decisive evidence** that students can accurately assess their own learning, and in fact suggests that several factors may influence students' self-evaluations, including learners' experience with a particular subject matter, self-esteem, and other psychological factors. Many who have researched self-assessment extensively in other fields of higher education maintain that frequently-observed tendencies for students to over- or under-assess argue **against** applications of self-assessment which are **not** formative in nature, and moreover, that students should be **trained** to assess their abilities in any given subject, in order to ensure the efficacy of the self-assessment instrument.

Second, research on self-assessment in foreign language pedagogy has been narrowly focused, for the most part, on concurrent or criterion validity, and has produced generally unimpressive results. Two-thirds of the 39 empirical studies published set out to establish the concurrent validity of a self-assessment instrument -- by comparing self-assessment results with either: a) the results of a previously-established, 'objective' test, b) a teacher's ratings of a student on the same scale, or c) a final course grade. Pearson Product-Moment correlations are the most commonly reported statistics used to measure concurrent validity, and reported correlations range from -.047 to .82, with most clustering between .3 and .6. Many of those studies which report a wide range of correlations between subtests, tend to focus their discussion of results only on the highest correlations and ignore the less impressive ones. More importantly, nearly everyone writing about an empirical investigation of self-assessment seems to have a different interpretation of what correlation levels actually indicate. According to LeBlanc &

Painchaud's earlier articles (1980, 1981, 1982, and others), a correlation of .49 is "good evidence" that students can self-assess, while Janssen-van Dieten (1989) calls correlations ranging from .29 to .69 "too low". Likewise, Wesche, Morrison, Ready, and Pawley (1990) deem a correlation of .58 "quite low", while Krausert (1991) argues that her correlations of .36 to .54 are "high". LeBlanc & Painchaud (1985) did report genuinely high correlations of .80 & .82, but these results stand alone and have not been replicated since.

Finally, because of the prevalent research objective of establishing concurrent validity, questions pertaining to **what** self-assessment is actually measuring, what cognitive processes underlie self-assessment, and how social and psychological factors may affect learners' self-assessments have only recently received any significant attention. Of the few studies which have addressed the question of other factors' influence on self-assessment, three have provided some evidence that self-esteem may play a role in how students evaluate their abilities (Anderson, 1982, Wesche, Morrison, Ready, and Pawley 1990, and Ready, in Press). Several others have claimed that question wording (metalinguistic versus situational and negatively versus positively-worded questions) may also affect self-assessment results (LeBlanc & Painchaud 1985, Evers 1981, and Bachman & Palmer, 1989).

Additionally Peirce, Swain & Hart, 1993 found some evidence of what they called the 'benchmark effect' on self-assessment; that is, students' self-assessment results using a situational (task-based) benchmark produce higher correlations with the criterion measurement than their self-assessments referencing a more global (peer-based) benchmark. Heilenman, 1990, found evidence in students' self-assessments of two 'response effects': acquiescence (a tendency to respond positively to any question), and social desirability (a tendency to respond so as to

appear to conform to perceived social values). She further ascertained that certain individual factors, including classroom experience and course currently enrolled in, grade, amount of inconsistent responses, experience in a French-speaking environment, and question or scale wording may contribute to the variance in self-assessment results. These two studies represent a promising beginning for much-needed, in-depth investigations into cognitive and social influences on language learners' perception and presentation of their own linguistic abilities.

The present study was specifically designed to explore the question of what cognitive processes and social-situational influences underlie students' assessment of their foreign language proficiency. The fundamental objective was to investigate the **process** of self-assessment through an interpretative examination of that process itself. That is, to focus on what students do when they self-assess, rather than on how well their overall score matches with some other overall measurement.

For this study, two types of verbal reports were used: think-aloud protocols and an immediate retrospective interview. The task in this case was that of responding to a self-assessment questionnaire consisting of sample questions from several previously-published reports on self-assessment. This questionnaire was administered to 28 learners of French from four distinctly different course levels. Summary descriptions of the subject pools may be found in Appendix I. Think-aloud protocols and retrospective reports in a semi-structured interview format were used to elicit subjects' thoughts both during and after the self-assessment process. Resultant data were analyzed for indications of: 1) students' orientation to the self-assessment task; 2) their interpretations of questions and rating scales; 3) possible influence of course levels

and experience with French or other languages; and 4) students' basic level of comfort in speaking French.

Questionnaire Development

The questionnaire used for this study consisted of 26 items covering the so-called 'four skills': listening, speaking, reading, and writing. All questions were taken directly from previously-established questionnaires (Anderson, 1982; Clark, 1981; Ferguson, 1978; Heilenman, 1990; LeBlanc & Painchaud 1982, & 1985; Oskarsson, 1980), with the goal of putting together a manageable number of questions which would be representative of the more common approaches to self-assessment, and would be sensitive to the different levels of language proficiency. There were four questions on writing, six on reading, seven on speaking, and nine on listening comprehension. Questions in each skill area ranged from very general wording: *'My understanding of what I read is:'* to more specific, situational wording: *'If I try to read a short newspaper article in French without a dictionary, I can get a general idea of what is going on.'* Questions also varied in terms of their difficulty levels (as defined by the author of the source questionnaires) e.g., *'If someone addresses me in French, I can understand the gist of what I am being told,'* versus *'I can understand discussions in French just as well as those in my mother tongue.'* A 5-point scale was indicated directly under each question, and questions were randomly ordered, rather than grouped by skill or by level of difficulty. Subjects were asked to verbalize out-loud everything that came to mind as they worked through the questions and decided on their answers. When they had finished the questionnaire, they were interviewed about specific answers they had given, as well as about some more general issues concerning the

self-assessment process they had just completed. Subjects' answers for both the think-aloud task and the follow-up interview were audio-taped, and later transcribed for analysis.

Analysis: When the data were analyzed, six principal categories of factors influencing the self-assessment process were identified. They are: Question Interpretation, Language Learning Background and Experience, Reference Points, Strategies for Completing a Questionnaire, Level of Certainty towards Answers, and Level of Self Confidence. Due to time limitations for this presentation, I will discuss only the issue of Reference Points today.

Researchers in Psychology have identified at least four possible reference points, which subjects may employ when they self-assess (see Higgins, Strauman, & Klein 1990). They are:

1) Social Category: A factual standard defined by the 'average' performance or attributes of the members of some social category or group -- in this case, subjects' current or past language learning colleagues.

2) Meaningful Other: A factual standard defined by the performance or attributes of another individual who is meaningful to the evaluator because of the relevance or appropriateness of that person's attributes for social comparison, or by reason of that person's emotional significance or importance to the evaluator. In this case, a meaningful other might be a fellow classmate, or a native speaker s/he once met, etc.

3) Autobiographical: A factual standard defined by the evaluator's own past performance or attributes. (These may be) a single instance, or a distribution of instances, recent or remote instances...(subjects) might compare different levels of achievement, or different amounts of change from levels of achievement.

4) Social Context: A factual standard defined by the performance or attributes of the immediate context of people to whom the evaluator is currently exposed (and notices). (Higgins, Strauman & Klein, 1986, p. 26)

Note that the fourth reference point, Social Context, is probably more relevant to psychological experiments which deliberately expose study participants to one or more people with certain characteristics, and then examine the influence of that exposure on subjects'

performance in the immediately-following experiment. In fact, no evidence of this reference point was found in the present study data.

The think-aloud and retrospective protocols for this study revealed many different examples of each of the three principal reference points described above. The first is ***Social category***. In the think-aloud data, several social categories were mentioned, the most popular being students' current classes. First, students generally provided rather imprecise descriptions of how they compared themselves to their classmates, as seen in examples 1, 2, and 3.

> **1) Subject 1-9 (F)[1]:** *...average is where the rest of the people maybe were in my class...*
>
> **2) Subject 2-1 (F)** *I think in comparison to my classmates, if I'm doing as well as my classmates, I'd say that's average. If I'm doing better, that's above average. So I, I assume every class is different. Like the class in general is the absolute scale, and I like sort of fit in, in terms of my ability. That's average.*
>
> **3) Subject 2-8 (F):** *I was thinking that, I seem to get better grades than the people I've talked to....*

Some subjects compared themselves with larger, more general groups, as evidenced in examples 4 & 5.

> **4) Subject 2-9 (Q12):** *...as far as how many people are my age who speak French, I'd say below average.*
>
> **5) Subject 4-3 (Q17):** *...as compared to other people who speak French, other Americans who speak French, maybe slightly above average.*

---

[1] Coding explanations: Subject 1-9 is the ninth subject in the first group of subjects (the lowest proficiency level); 'F' means the excerpt was part of a response to a follow-up question; 'Q12' means the excerpt was part of a response to question 12 on the questionnaire.

One has to wonder, however, how exactly students determine how well the members of these larger groups speak French. Finally, we see in example 6 that subject 1-3 made comparisons with a social group which he admitted may or may not actually exist.

> **6) Subject 1-3 (F):** *...those students who are studying at my level, and have basically the same exposure as I have. It's probably unlikely that somebody had as much proximity to a French-speaking environment as I did, growing up, so that's a little bit unfair...*

Several subjects mentioned the wider social category of Cornell University where they were studying. Such references were not substantiated with direct evidence about the quality of Cornell students, but rather seem to be based on Cornell's academic reputation. Students who see themselves as part of that community seem to feel that they must accordingly raise their standards for comparison, as illustrated by example 7:

> **7) Subject 2-2:** *I was thinking about that ... when I was about to say slightly below average on number 17, and I said, well, I'm at Cornell here, and you know, I mean, you're talking about below average. I mean, am I below average? ... I would say that average I would define is mediocre, meaning it's less than what I would want, to be average in anything. And ... like when I have a 'B' on a test, my friend and I have a joke, when I get back a 'B' on a test, we call it 'B-diocre.' I mean, it's not a bad grade, it just seems so average. So I'd say I'd describe average as not less than acceptable, but less than desirable. ... This is my first semester here.... (F - Q17)*

Clearly, this subject, a freshman, is influenced by his feeling of what it means to be studying at Cornell. He was about to say slightly below average, but then remembered where he was, and 'boosted' his evaluation of himself.

With regards to the ***Meaningful Other*** category, there were many fewer allusions in the think-aloud data to specific individuals who might be seen as Meaningful Other reference points; those who were mentioned may have been relevant because of their perceived abilities (or lack

BEST COPY AVAILABLE

of abilities), because of the level of their participation in the class, or because of their similar experiences, as seen in examples 8-11:

8) Subject 2-6: *Average is, you take the best person in the class, and the worst person in the class, and you say, I can relate more to this person than that person. (F)*

9) Subject 1-7: *In my class there are a couple of people who already know French and are taking it anyway... (F)*

10) Subject 2-4: *There's some people who ... whether or not they have more of a grasp of the language than I do, they were willing to speak... (F)*

11) Subject 3-2: *I think of the people in my class who've had similar experiences to me. For example, the girl who lived in Switzerland. I compare myself to her. (F)*

Subjects at all levels also referred to native speakers when making comparative evaluations of their language proficiency. These were not necessarily identified as specific individuals; rather, they seem to be either a perceived ideal native speaker, or the general community of native speakers, as illustrated in examples 12 and 13.

12) Subject 2-3: *Given a general kid who's grown up with French as his mother tongue, (I'm) way below average. (Q12)*

13) Subject 1-5: *I thought it meant average with respect to the French community, and then, well I was way below average... (F)*

This native speaker reference point may be based in experience for students who have lived in a Francophone country; for others, it is more likely a projection of their impressions of what a native French speaker could do, based on their knowledge of what they can do with their own native language.

Finally, the think-aloud data show evidence of numerous ways in which ***Autobiographical*** reference points may be expressed. First, subjects may compare their abilities in French with their abilities in other foreign languages, as we can see in example 14, where the student makes a comparison between his French and Spanish skills, or they compare their French with their native language, as illustrated in example 15, where the subject expresses her belief that her good writing abilities in English would carry over into French.

14) Subject 1-2: *I can give a presentation in Spanish, but in French, I don't think I could do that... (Q2)*

15) Subject 2-5: *I'm sort of proud of my writing abilities in English, so I think that maybe that would carry over to another language, and then I if I feel I can communicate ideas well in my own language, then I might be able to communicate ideas well in another language also, so ... that's probably why I feel slightly above average. (F)*

Subjects also make reference to their grades, for example, in 16 and 17:

16) Subject 1-10: *...I mean, average would be just like getting by, kind of barely. It probably depends on grades and understanding. (F)*

17) Subject 2-8: *I seem to get better grades than the people that I've talked to ... I mostly get B+s and A-s on what I'm writing, so I figured that would probably be slightly above average.*

In addition to mentioning grades, subjects refer to the amount of time they've spent studying French as well as to their goals and progress. They may also speak of past experiences, from which they try to determine their average level of proficiency, as evidenced below in examples 18-20:

18) Subject 3-1: *I've been speaking it for several years... // I'm finding now that I'm not using the dictionary as much as I used to ...I actually find that I'm more comfortable now... (Q12) // (F)*

19) Subject 4-4: *...it's a lot easier for me than it used to be. (Q4) // ...average compared to ... No, slightly above average compared to where I was last year, but not, I still feel like I have a lot to learn, especially about managing registers, and making my writing more stylistically interesting. (Q21)*

20) Subject 1-6: *...OK, my general ability to speak is, I would run through examples, um, of when I was speaking, thinking of how many times have I been corrected, how many times do I do this? ... I just rely on past experiences, but it wasn't a specific experience.... (F)*

Within the category of Autobiographical reference points, must also be included references to the amount of **effort** subjects perceive a particular task as requiring, as well as their perceptions of their abilities to get around in a Francophone country, whether they have actually had that experience or not. In response to the follow-up question, "**How do you determine what average is?**", students gave a variety of answers illustrating this reference point, represented in examples 21-24:

21) Subject 1-4: *I think it's how I speak. Like sort of easy. Way above average would be I wouldn't have to have a dictionary.*

22) Subject 2-8: *...if I think I can do something at least half the time, or if I can be correct half the time, I would say I was average at speaking French. I'm pretty much basing it on my own abilities.*

23) Subject 2-10: *When ... I have to sit and think about it, or if I can do just do it decently, then it's average. But if I can pass it off without having to think about it too much, then it's above average.*

24) Subject 3-3: *I guess average would be like functional. For instance if I said that my speaking ability is average, I wouldn't be like lost, on the streets, and I can get along, and ask directions, and I can talk to people. You know, very colloquially, or something.*

Some subjects, when trying to determine their proficiency in one language skill, compare their abilities in another skill, as in examples 25 and 26 in which subjects compare their written and spoken French:

25) Subject 1-1: *My written French is more correct than the French I speak, so I guess I somewhat agree (D) that my written French is correct. (Q25)*

26) Subject 2-5: *I would say that it's (writing ability) slightly above average, because I feel that I can write better than speak. (Q21)*

Interestingly, for the five students who made specific comparisons to other skills, the average scores for each skill did not reflect their general assessments. Though subject 1-1 felt she could write better than she could speak, her average speaking score was 3.3, while her average writing score was 2.4. Subject 2-5, who also felt she writes better than she speaks, rated her writing and speaking nearly identically: 2.8, and 3.0, respectively. These discrepancies may be a function of the limited number of questions for each skill on this experimental questionnaire, or it may be that because of varying orientations to the questionnaire (due to different interpretations and background, etc.), the questions do not effectively measure specific skills. Nevertheless, it is important to keep in mind that subjects may have a **general** autobiographical assessment of their abilities which is not necessarily borne out when looking at their responses to **individual** questions.

The fact that subjects use so many types of reference points is an important issue for self-assessment, and should impel us to seriously question whether or not comparisons between individual students' self-assessment results can give a true picture of relative language proficiency. Further complicating matters is evidence from this study that reference points often

overlap. That is, subjects may consider their abilities compared to people in their class as above average, but compared to native speakers as below average, and then take the 'average' of those two assessments as their final assessment. Moreover, subjects may consider several reference points for different questions, as evidenced by the responses of Subject 3-2 to four separate questions. In response to the question, "**In general my ability to speak is**", this subject answered using two different social category reference points, and chose an answer between the two:

> **27) Social Category:** *...Compared to the other people I grew up learning French with? It's way above average. Compared to people who've spent a year abroad, I'd say it's average, or even below average.* (Answered Average)

Later, in response to the question, **"My understanding of what I read is" (Q17),** he used an autobiographical reference point, focusing on the difficulty of the task:

> **28) Autobiographical -- Difficulty:** *...I can get the gist of what it is, so, average, slightly above average. There's always things that I don't understand.* (Answered D - slightly above average)

He used additional autobiographical reference points for a later question, referring this time to his own goals, and to his skills in English, as well as to the difficulty of the task, with his response to the question, **"My ability to write what I want is:" (Q21)**

> **29) Autobiographical -- Other language, Goals, Difficulty:** *I can never get across exactly what I want. So, I can never do that even in English, but I'll say average.* (Answered C)

Finally, in response to the follow-up question, **"How do you determine what average is for you?",** he mentioned several different reference points: the social category of Cornell (he is also

BEST COPY AVAILABLE

a Freshman), a specific classmate, his previous French class, and an autobiographical reference point to describe how well he can function in class, as may be seen in example 30.

> **30) Social Category -- Cornell; Meaningful Other -- Specific Classmate:** *Well, the average has been boosted a bit since I'm at Cornell, so the average should at least be the average of the students in the last French class I was in. ... the girl who lived in Switzerland, I compare myself to her...*
>
> **Social Category -- Class; Autobiographical -- Functionality:** *...I think normally, I can keep up, and I'm above average, maybe I make more mistakes, but I feel real comfortable, and I can at least speak, and making mistakes is part of it.*

Those researching self-assessment generally ask the question of whether students **can** evaluate their own language abilities. The preceding discussion of Reference Points should make it clear that self-assessments not only entail evaluation of one's **own** language proficiency; but involve judgment of the linguistic abilities of **other** people as well. This raises some serious, related questions: How well do students know what their classmates can do with the second language? On what grounds do they determine that someone is the best or worst student in the class, in order to make comparisons? Do students in fact have a realistic way of measuring their own proficiency? Excerpts from the follow-up interviews of two subjects may provide some preliminary (though not encouraging) answers. I asked them, **"How do you determine what average means?"** and had the following responses:

> **Subject P1-1** said, *Average? How would I define it? How well my classmates are doing.* // And on the next follow-up question she said, *I really don't know how well my classmates are doing.*
>
> **Subject 4-4 said,** *...I'm not really sure where I am, partly because I've read very little writing from other people. I don't really know how good they are, but they're teaching French, so I assume they're pretty good...*

In conclusion, the think-aloud protocols in this study have shown that subjects make use of a wide variety of benchmarks or reference points when evaluating their second language abilities. Substantial evidence of three reference points in particular, Social Category, Meaningful Other, and Autobiographical, were found in the present study data. The Social Category reference point comprised references to classmates, to larger, more generally-defined groups, such as 'all the people my age who speak French', or to the immediate context of a university with a good academic reputation. Meaningful Others identified in the data included specific classmates, and native speakers. Autobiographical references were made to students' experiences with other languages, to their grades, to their goals and sense of progress or enjoyment, and to the amount of effort they feel they expend in the learning process.

Two important issues emerge from this discussion of self-assessment reference points. First, reference points are numerous and varied, and subjects use them in equally diverse ways. One subject may use several different reference points throughout the questionnaire, or in fact, in response to a single question. Second, with the exception of the autobiographical reference point, students rely on their perceptions not only of their own, but of others' linguistic abilities to make their judgments. It is doubtful that the majority of students has a well-grounded understanding of anyone else's language proficiency, which further calls into question the possibility that second language students can effectively assess their own abilities. It might be suggested that the reference points could be specified clearly enough before a student begins the self-assessment process, so that problems with differing comparisons could be alleviated. Evidence from the analysis of the varying influences of language learning background and experience, and question interpretation, also examined in this study, however, compel me to state

that it seems highly unlikely that designating a particular benchmark will ensure that students will all interpret that reference point in like fashion.

For some, the overriding question concerning self-assessment is one of efficacy: how well can students evaluate their own second language abilities? By contrast, the present study has taken as a starting point the fact that previous results pertaining to efficacy are unconvincing, and in fact suggest there is more to self-assessment than students making reasonable judgments about second language proficiency. What does it really mean to self-assess? What is self-assessment really measuring? Answers to both these questions lie in the individuality of the subjects. Self-assessment measures individuals' responses to the questionnaire, based on their individual interpretations of questions and rating scales, and influenced by individual experiences and language learning backgrounds, as well as individually-determined strategies for approaching the self-assessment task, individually-defined points of comparison (which we saw in some detail today), and individual levels of self-confidence, both with regard to foreign language abilities, and to answers on the questionnaire. This individuality, in many ways, and at many levels, compromises both the validity and the potential efficacy of self-assessment as a tool for measuring second language proficiency. It may be concluded, therefore, that trying to fit that individuality onto a three-, five-, ten- or whatever-point scale is, for many practical purposes, an impossible task. In the case of applications such as placement testing, which necessarily entail comparisons and rankings of large numbers of students, self-assessment cannot be a valid summative measurement tool.

The evidence from this study should not rule out its use in formative situations, however. Discussion of one's abilities and progress over the course of a learning program has been shown

to be quite useful in many other fields of higher education. It is hoped that the present study will contribute to discussions of self-assessment in foreign language education, in terms of developing a better understanding of the processes underlying self-assessment, and of the ways it might be used to effectively enhance students' language learning experiences.

Student Self-Assessment of Language Proficiency: Perceptions of Self and Others

Christine E.B. Moritz
The University of South Carolina
AAAL Chicago -- March 24, 1996

Appendix I -- Subject Descriptions

Group-1: Ten students completing their second semester of beginning French, with little to no experience with French prior to beginning French at Cornell University. One subject had lived in Canada, and had chosen to skip the first-semester class. Another subject had had four years of high school French, but placed into French 122 upon entering the University; three students had spent some time abroad (up to 5 weeks). Two considered themselves bilingual (English/Spanish & English/Russian); only one subject had not studied any language other than French.

Group-2: Ten students completing a 200-level French course, with an average of four years of high school French, and two semesters of college French. One had spent a good deal of time in Paris, where he has a French girlfriend, but the remainder had not had any experience living in a Francophone country, though two had relatives who spoke French. Two bilingual (English/German and English/Cambodian); four had not studied another foreign language besides French.

Group-3: Four students completing a 300-level course on French cinema, with an average of four to six years in junior high and high school, and all had spent from six months to a year living in France. One had studied French for only one year in high school, but had lived in Geneva, Switzerland, for more than three years. All had taken just one or two semesters of university-level French, and all reported little or no experience with languages other than French.

Group-4: Four graduate students with between nine and 21 years experience learning and speaking French. All employed as teaching assistants in French at Cornell University. All had lived in a Francophone country, from four months to two and a half years. All had studied Spanish, from two to 13 years, and two had lived in Spanish-speaking countries. None raised bilingually.

Appendix II -- Selected References

Anderson, P. 1982. Self-Esteem in the Foreign Language: A Preliminary Investigation. *Foreign Language Annals,* 15(2).

Bachman, L.F. and Palmer, A.S. 1989. The Construct Validation of Self-Ratings of Communicative Language Ability. *Language Testing,* 6(1), 14-29.

Blanche, P. and Merino, B.J. 1989. Self-Assessment and Foreign Language Skills: Implications for Teachers and Researchers. *Language Learning,* 39(3), 313-340.

Cohen, A.D. 1982. Introspecting about Second Language Learning. *Studia Anglica Posnaniensia: An International Review of English Studies,* 15, 139-146.

Ericsson, K.A. and Simon, H.A. 1993. Protocol Analysis: Verbal Reports as Data. Cambridge, MA: MIT Press.

Falchikov, N. and Boud, D.J. 1989. Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research,* 59(4), 395-430.

Heilenman, L.K. 1991. Self-Assessment and Placement: A Review of the Issues. in. R.V. Teschner, ed., *Assessing Foreign Language Proficiency of Undergraduates.* Boston: Heinle & Heinle, pp. 93-114.

Heilenman, L.K. 1990. Self-Assessment of Second Language Ability: the Role of Response Effects. *Language Testing,* 7(2), 174-201.

Higgins, E.T., Strauman, T., and Klein, R. 1986. Standards and the Process of Self-Evaluation: Multiple Affects from Multiple Stages. in Sorrentino, R. and Higgins, E.T. eds., *Handbook of Motivation: Foundations of Social Behavior.* New York: The Guilford Press. pp. 23-59.

Holec, H. 1988. Autonomy and Self-Directed Learning: Present Fields of Application. Strasbourg: Council of Europe: Council for Cultural Cooperation.

Holec, H. 1980. Learner Training: Meeting needs in self-directed learning. in H.B. Altman, and C.V. James, eds. *Foreign Language Teaching: Meeting Individual Needs.* Papers from the First Pergamon Institute of English Seminar, Oxford, 1979. Pergamon Press: Oxford, England.

Janssen-van Dieten, A.M. 1989. The development of a test of Dutch as a second language: the validity of self-assessment by inexperienced subjects. *Language Testing,* 6(1), 30-46.

Krausert, S.R. 1991. Determining the Usefulness of Self-Assessment of Foreign Language Skills: Post-Secondary ESL Students' Placement Contribution. Ph.D. Dissertation, University of Southern California.

LeBlanc, R. et Painchaud, G. 1985. Self-Assessment as a Second Language Placement Instrument. *TESOL Quarterly,* 19(4), 673-687.

Lewkowicz, J.A. and Moon, J. 1985. Evaluation, a way of involving the learner. in J.C. Alderson, ed. *Lancaster Practical Papers in English Language Education, vol. 6: Evaluation.* Oxford: Pergamon Press. pp. 45-80.

Moritz, C. 1995. Self-Assessment of Foreign Language Proficiency: A Critical Analysis of Issues and a Study of Cogniiive Orientations of French Learners. Ph.D. dissertation. Cornell University.

Oskarsson, M. 1989. Self-Assessment of Language Proficiency: Rationale and Applications. *Language Testing*, 6(1), 1-13.

Oskarsson, M. 1988. Self-Assessment of Communicative Proficiency. in Trim, J. ed. *Evaluation and testing in the learning and teaching of languages for communication.* Strasbourg: Council of Europe, Council for Cultural Cooperation.

Oskarsson, M. 1984. Self-Assessment of Foreign Language Skills: a survey of research and development work. Strasbourg: Council of Europe, Council for Cultural Cooperation.

Peirce, B., Swain, M., and Hart, D. 1993. Self-Assessment, French Immersion, and Locus of Control. *Applied Linguistics*, 14(1), 25-42.

Ready, D. In press. The role and limitations of self-assessment in testing and research. *Proceedings LTRC 1991.* Princeton, N.J.: Educational Testing Service.

Wesche, M.B., Morrison, F., Ready, D. and Pawley, C. 1990. French Immersion: Postsecondary Consequences for Individuals and Universities. *Canadian Modern Language Review*, 46(3), 430-451.

Wesche, M., Paribakht, T.S., and Ready, D. 1993. A Comparative Study of Four Placement Instruments. *ERIC Document #ED364064.*